Applied Energy 228 (2018) 1111–1121

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

Wind power field reconstruction from a reduced set of representative measuring points

S. Salcedo-Sanz[a,*], R. García-Herrera[b,c], C. Camacho-Gómez[a], A. Aybar-Ruíz[a], E. Alexandre[a]

[a] Department of Signal Processing and Communications, Universidad de Alcalá, Madrid, Spain
[b] Department of Earth Physics and Astrophysics, Universidad Complutense de Madrid, Madrid, Spain
[c] IGEO, Instituto de Geociencias (UCM/CSIC), Spain

HIGHLIGHTS

- We deal with a problem of representative measuring points selection for long-term wind power analysis.
- We select the best set of N measuring points which produce the least wind power reconstruction error.
- We apply a novel meta-heuristic for solving the problem.
- The Analogue Method is used as wind power field reconstruction algorithm.
- Results for Europe with ERA-Interim reanalysis data are reported.

ARTICLE INFO

Keywords:
Wind power field reconstruction
Wind resource analysis
Representative points
Coral Reefs Optimization algorithm
Analogue method

ABSTRACT

In this paper we deal with a problem of representative measuring points selection for long-term wind power analysis. It has direct applications such as wind farm prospective location or long-term power generation prediction in wind-based energy facilities. The problem's objective is to select the best set of *N* measuring points (i.e. *N* representative points), in such a way that a wind power error reconstruction measure is minimized, considering a monthly average wind power field. In order to solve this problem, we use a novel meta-heuristic algorithm, the Coral Reefs Optimization with Substrate Layer, which is an evolutionary-type method able to combine different search procedures within a single population. The CRO-SL is hybridized with the Analogue Method as wind power reconstruction method, to identify the most representative points for the wind field. The proposed approach has been tested in the reconstruction of monthly average wind power fields in Europe, from reanalysis data (ERA-Interim reanalysis). The method exhibits strong performance as evidenced from the experiments carried out. The solutions obtained show that the more significant measuring points are mainly located over the Atlantic ocean, which is consistent with the wind speed climatology of the Northern hemisphere mid-latitudes. We have also analyzed the set of least representative points to reconstruct the wind power field (less informative points for whole reconstruction of the field), obtaining points mainly located at the North of Scandinavia (which may be associated with the circumpolar circulation), and some points in the Eastern Mediterranean, which seem to be related to the Etesian winds. Reconstructions at seasonal scales show similar results, which provides confidence on the robustness of the proposed method. The proposed methodology can be further applied to alternative energy-related problems, such as the selection of critical energy infra-structures or the selection of critical points for climate change studies, among others.

1. Introduction

Long-term wind speed and power prediction and variability studies are very useful in different wind farm tasks, such as wind farm location and prospective works, maintenance planning, financial estimates or long-term electricity generation prediction [1]. However this, long-term wind power production analysis has not received much attention in the literature. Several studies are focused on variability studies of monthly and seasonal wind speed and related power. In [1], ERA-Interim reanalysis data are used to model part of the long-term variability of the wind energy resource in central Europe by reconstructing surface wind distributions at different long-term scales. In [2] an adaptive

* Corresponding author.
E-mail address: sancho.salcedo@uah.es (S. Salcedo-Sanz).

https://doi.org/10.1016/j.apenergy.2018.07.003
Received 12 April 2018; Received in revised form 12 June 2018; Accepted 1 July 2018
Available online 07 July 2018
0306-2619/ © 2018 Elsevier Ltd. All rights reserved.

general regression neural network is used to carry out spatial regression of monthly wind speed in the Alps considering NCEP/NCAR reanalysis data. In [3] an analysis of the variability of monthly average wind speed in Hungary is carried out. The methodology uses a Weibull distribution to model the wind speed at different altitudes, showing that this distribution seems optimal at monthly scale. In [4] a study of multi-decadal variability of daily wind speed is carried out using NCEP/NCAR reanalysis data. In this case, an Evolutionary Computation algorithm was used to obtain different classes of wind speed, and study their inter-decadal variability. This work was extended in [5], where the long-term variability of wind power in Iberia was analyzed. Other papers are focused on the statistical evaluation of monthly, seasonal or annual wind speed features in a given zone. For example, in [6] a statistical analysis on the monthly wind characteristics, using measured wind speed data for a five years period in the Southern Egyptian desert is carried out. The results show that a 150 Mw wind farm could be successfully planned in the region to supply energy covering the necessities of the nearest cities. In [7] an analysis of monthly average wind speed data in Chad is performed using Weibull distributions. In [8] the authors try to identify whether statistically significant differences exist among the monthly wind average speed of the four major Brazilian states on wind energy productive capacity. They apply the Nested Gage Repeatability & Reproducibility, generally used on manufacturing quality management, to the wind speed series of the different States in order to identify major differences among them.

The analysis of the wind speed climatology in different regions of the World, including climate change, is the purpose of other studies. For example, in [9] an analysis of the South American long-term wind speed climatology is carried out, taking into account data from MERRA reanalysis and the effect of the main climate drivers in the zone, such as the Antarctic Oscillation (AAO) or El Nino/Southern Oscillation (ENSO). In [10] long-term variability of wind indices is discussed. In [11] a study of wind extremes for wind power generation in England is carried out with reanalysis data from the last 30 years. Studies of the climate change impact on wind energy always consider long-term analysis of wind speed data. For example, in [12] a review of climate change impacts in wind energy is carried out, with specific analysis on wind resource (long-term wind speed), operation and maintenance of facilities. A case study in the Scandinavian zone is analyzed. In [13] an analysis of five different Atmosphere-Ocean General Circulation Models to investigate potential changes in long-term wind speeds and energy density in northern Europe is carried out. In [14] an analysis of climate change in wind energy in Norway is presented. Different scenarios of climate change impact are supposed, and in all them it is shown that wind energy could help reduce energy deficit in Norway. Finally, in [15] an study on wind resource variability in the USA is carried out, by studying the output of coupled atmosphere-ocean general circulation models. The models' simulation show that the long-term wind resource in the USA regions of largest wind energy penetration will maintain the historical envelope of variability. Thus it seems that this renewable resource will be available in the next decades. However, it must be taken into account that wind is perhaps the meteorological variable more difficult to forecast, due to its extreme spatio-temporal variability. This limits the geographical representativeness of the wind measuring points, and therefore affects subsequent studies related to wind energy resource. On the other hand, the density of wind observation sites is much lower than those for other variables, such as temperature or precipitation, specially over the oceans, not only in Europe [16], but also in other parts of the world [17]. Thus, characterizing wind speed and power variability out of the observational points is usually much less reliable than for any other meteorological variable [18]. There have been different works dealing with uncertainty of wind speed resource analysis due to this limited data, for example [19] where Monte Carlo simulations for a long period of time by using limited data is suggested as methodology. Uncertainty of wind speed due to measurements limitation in different zones of the World have also been analyzed, such as in Caspian Sea [20], or in the Persian Gulf [21].

In this paper we analyze how wind speed and power fields can be reconstructed from a limited number of representative points able to capture the main features of the long-term variability. This problem is significant for wind speed prediction and wind power resource analysis, for different reasons: the reconstruction of wind speed series [22] is an important topic in wind speed prediction which is usually carried out for wind speed series using neighbor stations [23]. The study of long-term wind speed and variability is very relevant for wind resource analysis and evaluation, wind farms location and prospection problems [24]. In these applications, the selection of a representative subset of points to characterize the wind speed and power potential is a challenging problem. In fact, traditional optimization algorithms do not work properly due to the discrete nature of the encoding, and the black-box characteristic of the objective function (outcome of the Analogue reconstruction algorithm). From a different point of view, the solution of this problem should allow the identification of the best location of measuring points to accurately evaluate a given wind speed and power field. The selection of exemplar samples from a given data, points or items collection, in such a way that the selected exemplars accurately summarize the complete set of starting data is usually known as *Representative Selection* (RS) problem [25]. The RS problem from a field of wind speed measures may have alternative applications in renewable energy and meteorology problems, such as climate models downscaling [26], identifying the most representative subset of global climate models in terms of a given error measure [27], the selection of regional climate scenarios [28], choosing the most representative models for climate change studies [29], or optimizing the position of weather monitoring stations [30]. In all these cases, the idea is to keep the points which provide the maximum information for the problem at hand, while discarding those less informative points. In this paper we apply a meta-heuristic algorithm to solve a specific application of the RS problem related to wind power reconstruction. The so-called Coral Reefs Optimization algorithm with Substrate Layer (CRO-SL), is used as searching meta-heuristic in this case. It is an evolutionary-type algorithm able to combine different classes of search patterns within a single population. The basic CRO and the advanced CRO-SL have shown excellent performance in optimization problems related to renewable energy applications. For example, in [31] where the basic CRO is used for selecting the best set of features for a neural prediction system in a problem of wind speed prediction in wind farms. The CRO has also been applied to a problem of solar radiation down-scaling in Spain [32], and more recently for improving a neural network performance in a problem of solar radiation prediction in Australia [33]. The RS problem considered in this paper uses ERA-Interim reanalysis data in Europe to evaluate the optimal selection of points which lead to the best monthly-average reconstruction of wind power. The results obtained show the location of the points providing the highest information for wind speed and power reconstruction in Europe, when the Analogue Method (AM) is considered as reconstruction approach. In terms of novelty, this paper introduces a new methodology to select in an optimal way a subset of stations or measuring points. The use of a meta-heuristic approach provides robustness to the method, since the objective function can be changed to tackle very different specific applications, while maintaining the encoding and the searching approach. In this case, our proposal is a novel way of facing a problem of long-term wind power variability, which may be also useful for alternative problems with a similar structure, such as the analysis of other renewable energy resources in Europe.

The rest of the paper is structured in the following way: next section details the RS problem definition tackled as an optimization task. Section 3 summarizes the CRO-SL algorithm, with details on specific implementation of different search operators for the RS problem. Section 4 presents the experimental section of the paper, where we evaluate the performance of the approach in the RS problem. Finally, Section 5 closes the paper by giving some final conclusions and remarks.

2. Methodology and working data

This section presents the methodology taken into account for wind power field reconstruction based on the Analogue Method. Before, we summarize the working data necessary to apply the reconstruction algorithm, mainly data partition into train and test periods of (possible) different duration and definition of the optimal subset of measuring points to reconstruct the wind power field.

2.1. Working data

Let $F(\mathbf{s}, t)$ be a field of wind speed or power measurements, defined in a set of $|S|$ measuring points or stations $\mathbf{s}$, during an observation period $\hat{t}$. Such observation period can be split into a training period, t^T, and a validation or test period of duration t^V, in such a way that $\hat{t} = t^T + t^V$. Let χ be a given reconstruction algorithm for $F(\mathbf{s}, t)$. The reconstruction algorithm χ operates in a subset of measuring points $\mathbf{s}_N$, where N ($N < |S|$) represents the number of measuring points selected. Moreover, note that χ only uses the data in the training period to obtain the best reconstruction of the initial field $F(\mathbf{s}, t)$, in terms of an error measure $e(\mathbf{s}_N)$ evaluated in the test period. The problem we face in this paper consists of obtaining the best possible subset of N measuring points $\mathbf{s}_N^*$, which minimizes $e(\mathbf{s}_N^*)$ (usually e stands for a mean square or absolute error function). Note that the subset $\mathbf{s}_N^*$ stands for the N most representative measuring points for the wind speed field $F(\mathbf{s}, t)$, which allows the best reconstruction of the field in terms of the algorithm χ considered. As stated previously, in this paper we have chosen the well-known AM as the reconstruction algorithm χ.

2.2. The analogue method

The AM is based on the principle that two similar states of the atmosphere lead to similar local effects [34]. More specifically, two states of the atmosphere are considered as "analogues", when there is a resemblance between them, in terms of an analogy criterion and objective variables. Thus, the AM consists of searching for a certain number of past situations in a meteorological archive, in such a way that they present similar properties to that of a target situation for any chosen predictors or variables.

Different versions of the AM can be found in the literature, with a wide range of energy-related applications. In [35], an AM method has been applied to probabilistic solar power forecast in three solar farms in Italy. A comparison with a quantile regression algorithm and persistence ensembles has proven the goodness of the AM approach in this problem. In [36] an AM ensemble was applied to a problem of short-term wind power prediction. In this case an AM ensemble has been applied to the wind power production prediction of a wind farm in northern Sicily (Italy), comparing the performance with alternative algorithms such as quantile regression and numerical weather models. In [37] another AM ensemble approach is presented, with application in wind resource estimation. The paper analyzes the wind resource of different locations in Europe and USA by applying an AM ensemble.

Finally, note that very recently the AM has been mixed with meta-heuristics algorithms. In [38] a genetic algorithm has been used to tune the parameters of an AM. The accuracy of this evolutionary-AM approach has been shown in a case study of probabilistic precipitation forecast in Switzerland.

In this paper, we use the AM as the reconstruction algorithm χ. Given a subset of measuring points of stations $\mathbf{s}_N$, the AM process starts by obtaining the most similar situations in the past for the field $F(\mathbf{s}_N, t^V)$ (considering all the evaluation period). In other words, this is equivalent to, for each time $T \in t^V$, obtaining the most similar situation (or average of k most similar situations) in the past (training period), located in time $T^* \in t^T$, considering only the selected measuring points $\mathbf{s}_N$ (note that T^* depends on the T considered, i.e. $T^* = T^*(T)$). Then, the complete reconstruction of the field F is calculated by using the past

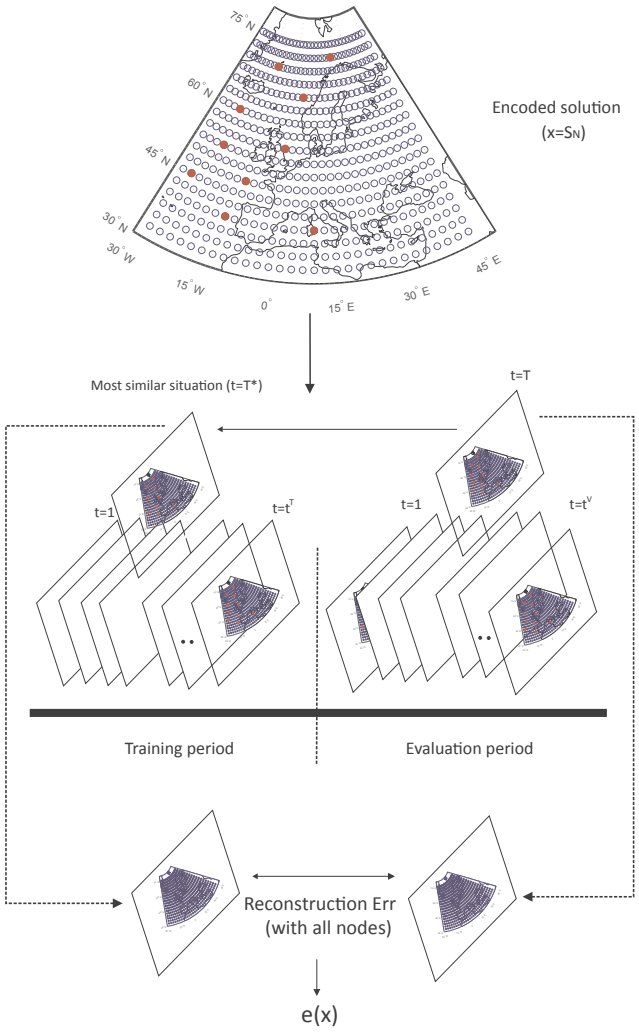

Fig. 1. AM calculation process. From a selection of measuring points to be evaluated (marked in red), the AM approach works by calculating the field reconstruction error for all the evaluation period, by looking for the most similar situation in the training period, considering only the information provided by the selected measuring points. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

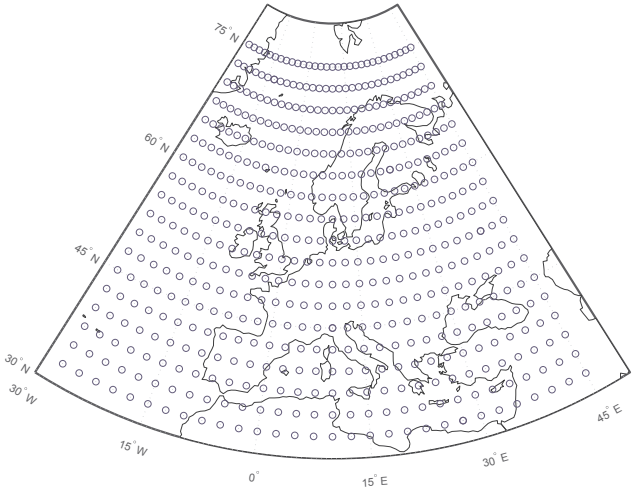

Fig. 2. Location of the measuring points in the ERA-Interim reanalysis nodes considered.

Table 1
Parameters of the optimization meta-heuristics compared in this paper: CRO-SL, HS, EA, PSO and SA.

	Parameters
CRO-SL	
Initialization	Reef size = 50 × 40 (2000 positions), $\rho_0 = 0.9$
External sexual reproduction	$F_b = 0.80$, $\mathscr{T} = 5$ substrates: HS, DE, 2Px, MPx, SM
Internal sexual reproduction	$1-Fb = 0.20$
Larvae setting	$\kappa = 3$
Asexual reproduction	$F_a = 0.05$
Depredation	$F_d = 0.15$, $P_d = 0.05$
Stop criterion	$k_{max} = 1000$ iterations
HS	
Harmony Memory	2000 harmonies
HMCR	0.8
PAR	0.3
Stop criterion	$k_{max} = 1000$ iterations
EA	
Population	2000 individuals
Selection	Tournament
Crossover	Two-points
Mutation	Uniform integer
Stop criterion	$k_{max} = 1000$ iterations
PSO	
Swarm	2000 particles
Learning rates	$\phi_1 = 0.5$, $\phi_2 = 0.5$
Maximum velocity	$V_{max} = 10$
Stop criterion	$k_{max} = 1000$ iterations
SA	
Temperature Change	$T_{n+1} = \frac{T_n}{0.01 \cdot (n+1)}$
Changes per temperature	2000
Mutation	Standard integer Mutation (SM)
Stop criterion	$k_{max} = 1000$ iterations

situation $F(\mathbf{s}, T^*)$ and the objective situation $F(\mathbf{s}, T)$, and a reconstruction error is obtained. The final function $e(\mathbf{s}_N)$ is calculated as the root mean square error (RMSE) in the field reconstruction:

$$e(\mathbf{s}_N) = \sqrt{\frac{1}{|S| \cdot t^V} \sum_{T=1}^{t^V} \sum_{k=1}^{|S|} (F(s_k, T^*) - F(s_k, T))^2} \quad (1)$$

Fig. 1 graphically shows the process for the AM application (χ field reconstruction algorithm), and the calculation of the error function (RMSE) associated with $\mathbf{s}_N$.

3. Optimization method: the CRO-SL algorithm

The Coral Reef Optimization algorithm (CRO) [39] (and further describe in [40]), is an evolutionary-type algorithm based on the behavior of the processes occurring in a coral reef. In [41], a new version of the algorithm (CRO-SL) was presented, where several *substrate layers* (specific parts of the population) were introduced. In this algorithm, each substrate layer may represent different processes (different models, operators, parameters, constraints, repairing functions, etc.). Specifically, in [42] a version of the CRO-SL algorithm was proposed, in which each substrate layer represents a different search procedure, leading to a co-evolution competitive algorithm. This version of the CRO-SL has been successfully tested in different applications and problems such as micro-grid design [43], vibration cancellation in buildings, both with passive models [44], and active models [45], or in the evaluation of novel non-linear search procedures [46]. This is also the CRO-SL algorithm used in this paper for optimizing the best subset of representative points for wind speed reconstruction.

Regarding the algorithm's encoding for the optimization problem at hand, we consider integer vectors as solutions, $\mathbf{x} \in \mathbb{N}$. Note that using this encoding the length of each individual is equal to N, and in this case it is necessary to control that there are not repeated measuring points in a given solution for it to be feasible. This encoding provides a compact version of the algorithm, and allows using some different searching procedures such as Harmony Search or Differential Evolution. Next subsection details the different substrate layers considered in the CRO-SL for this problem.

Table 2
Results in terms of the RMSE for the wind power field reconstruction (W/m^2) obtained in the Europe reanalysis data, by the CRO-SL, HS, EA, PSO and SA approaches.

	Best	Mean	Var
		CRO-SL	
$N = 5$	**0.27**	0.28	$3e^{-2}$
$N = 10$	**0.22**	0.23	$2e^{-2}$
$N = 15$	**0.19**	0.20	e^{-2}
$N = 20$	**0.18**	0.19	$2e^{-2}$
		HS	
$N = 5$	0.31	0.38	e^{-2}
$N = 10$	0.27	0.31	$2e^{-2}$
$N = 15$	0.25	0.28	e^{-2}
$N = 20$	0.23	0.27	$3e^{-2}$
		EA	
$N = 5$	0.28	0.29	$3e^{-3}$
$N = 10$	0.24	0.25	e^{-3}
$N = 15$	0.23	0.24	e^{-3}
$N = 20$	0.20	0.22	$2e^{-3}$
		PSO	
$N = 5$	0.29	0.32	$3e^{-3}$
$N = 10$	0.25	0.29	$2e^{-3}$
$N = 15$	0.24	0.28	$2e^{-3}$
$N = 20$	0.22	0.25	e^{-3}
		SA	
$N = 5$	0.30	0.31	$6e^{-3}$
$N = 10$	0.27	0.28	$3e^{-3}$
$N = 15$	0.24	0.27	$2e^{-3}$
$N = 20$	0.23	0.24	e^{-3}

Table 3
Best result by the CRO-SL approach compared with a randomly generated solution. Different statistical measures are reported: Root mean square error (RMSE), correlation coefficient (CC), bias (Bias), mean absolute error (MAE), normalized root mean square error (NRMSE) and scatter index (SI), corresponding to the ERA-Interim reanalysis data.

	RMSE	CC	Bias	MAE	NRMSE	SI
CRO-SL						
$N = 5$	0.27	0.54	$1.77e^{-5}$	0.12	$1.00e^{-3}$	$1.00e^{-3}$
$N = 10$	0.22	0.54	$2.64e^{-8}$	0.09	$8.23e^{-4}$	$8.23e^{-4}$
$N = 15$	0.19	0.55	$-2.03e^{-6}$	0.08	$7.39e^{-4}$	$7.38e^{-4}$
$N = 20$	0.18	0.55	$-1.05e^{-6}$	0.08	$6.96e^{-4}$	$6.96e^{-4}$
Random						
$N = 5$	0.49	0.50	$-3.00e^{-5}$	0.18	$1.8e^{-3}$	$1.8e^{-3}$
$N = 10$	0.50	0.50	$-3.03e^{-5}$	0.19	$1.9e^{-3}$	$1.8e^{-3}$
$N = 15$	0.49	0.50	$-4.02e^{-5}$	0.18	$1.8e^{-3}$	1.8^{-3}
$N = 20$	0.50	0.50	$-2.64e^{-5}$	0.18	$1.86e^{-3}$	$1.86e^{-3}$

3.1. Substrates considered in the CRO-SL

The considered substrates for solving this problem of selection of the best measuring points for optimal reconstruction of a wind power field

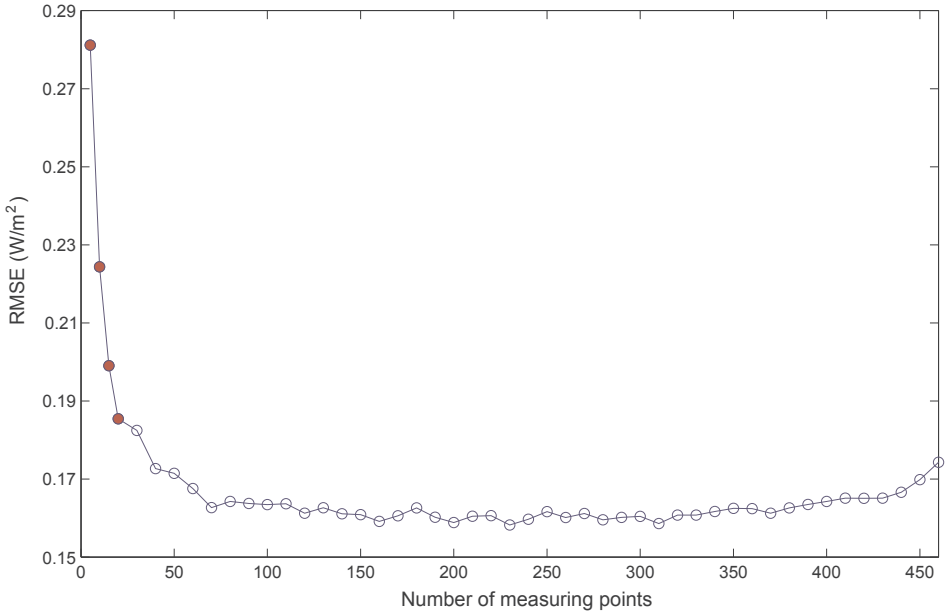

Fig. 3. RMSE (W/m^2) obtained with the CRO-SL approach for different values of N.

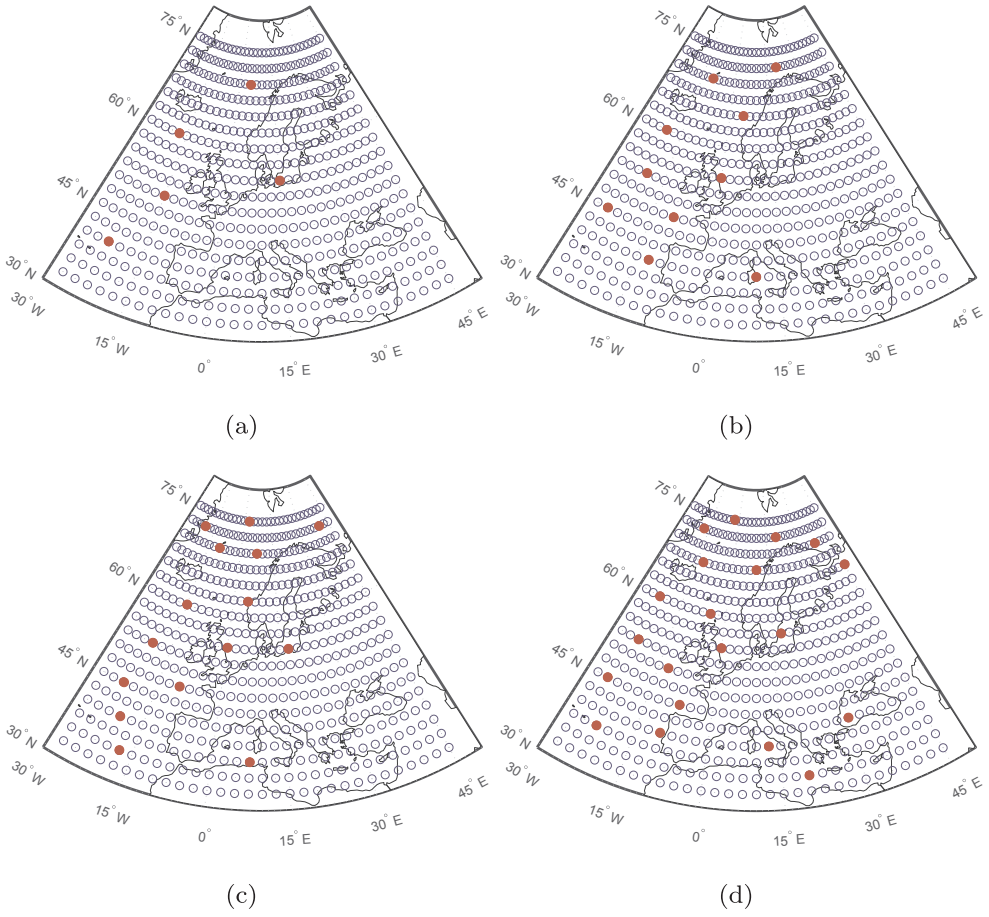

Fig. 4. Best solution found by the CRO-SL (red points stand for the selected representative nodes); (a) $N = 5$; (b) $N = 10$; (c) $N = 15$ and (d) $N = 20$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

are detailed below. Note that there are general purpose substrates, such as Differential Evolution or Harmony Search-based, and other specific substrates with crossovers and mutations adapted to the chosen encoding. A total of 5 substrates will be described and evaluated later in the experimental section.

- **Differential Evolution-based operator** (DE): This operator is based on the evolutionary algorithm with that name [47], a method with powerful global search capabilities. DE introduces a differential mechanism for exploring the search space. Hence, new larvae are generated by perturbing the population members using vector

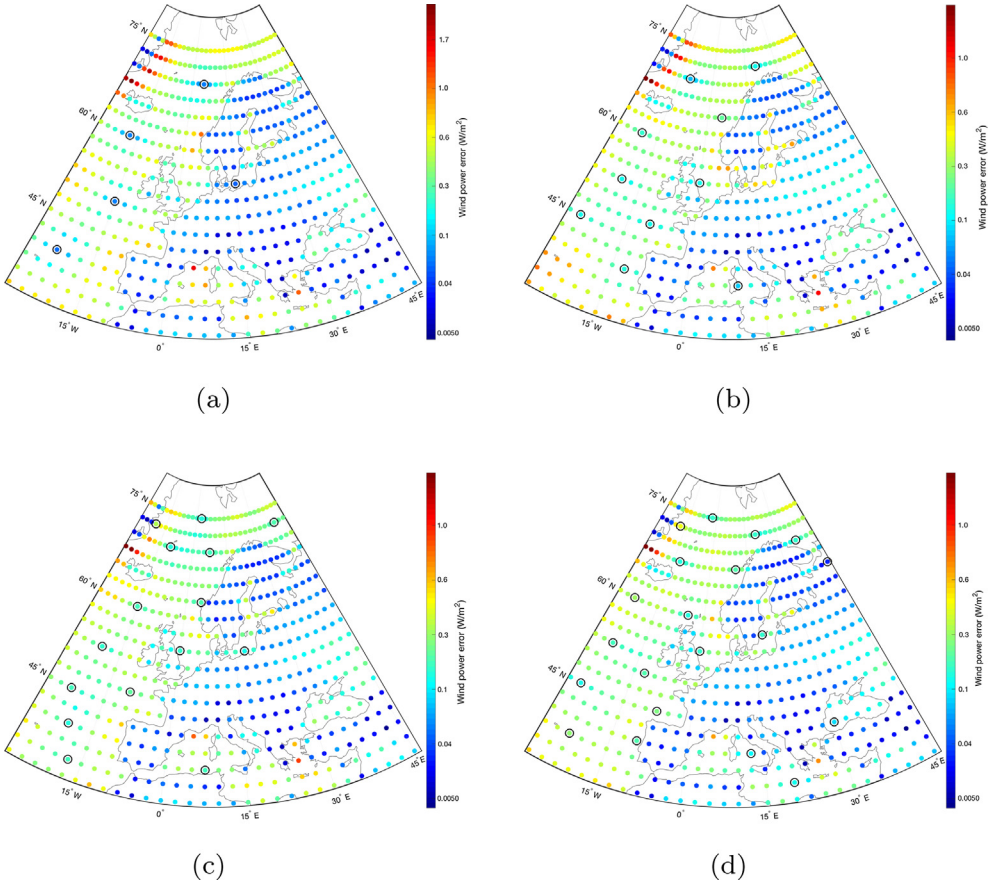

Fig. 5. Reconstruction error of the wind power field with the Analogue method, for different number of selected representative points; (a) $N = 5$; (b) $N = 10$; (c) $N = 15$ and (d) $N = 20$.

differences of individuals. Perturbations are introduced by applying the rule $x'_i = x_i^1 + F(x_i^2 - x_i^3)$ for each encoded parameter on a random basis, where x' corresponds to the output larva, x^t are the considered parents (chosen uniformly among the population), and F determines the evolution factor weighting the perturbation amplitude.

- **Harmony Search-based operator** (HS): Harmony Search [48] is a population based MH that mimics the improvisation of a music orchestra while its composing a melody. This method integrates concepts such as harmony aesthetics or note pitch as an analogy for the optimization process, resulting in a good exploratory algorithm. HS controls how new larvae are generated in one of the following ways: (i) with a probability HMCR $\in [0, 1]$ (Harmony Memory Considering Rate), the value of a component of the new larva is drawn uniformly from the same values of the component in the other corals. (ii) with a probability PAR $\in [0, 1]$ (Pitch Adjusting Rate), subtle adjustments are applied to the values of the current larva, replaced with any of its neighboring values (upper or lower, with equal probability).
- **Two points crossover** (2Px): 2PX [49] is considered one of the standard recombination operators in evolutionary algorithms. In the standard version of the operator, two parents from the reef population are provided as input. A recombination operation from two larvae is carried out by randomly choosing two crossover points, interchanging then each part of the corals between those points.
- **Multi-points crossover** (MPx): Similar to the 2PX, but in this case the recombination between the parents is carried out considering a high number of crossover points (M), and a binary template which indicates whether each part of one parent is interchanged with the corresponding of the other parent.
- **Standard integer Mutation** (SM): This operator consists of a standard mutation in integer-based encodings. It consists of mutating an element of a coral with another valid value (different from the previous one). Note that the SM operator links a given coral (possible solution) to a neighborhood of solutions which can be reached by means of a single change is an element of the coral.

4. Experimental evaluation

This section describes the different experiments and the results obtained. First, we describe the data used and the algorithms considered for comparison. Then, the main results obtained are summarized and discussed, in terms of the long-term wind power reconstruction, the distribution of the best subset of measuring points, a seasonal study and an experiment to evaluate the worst points for the field reconstruction. We finish the experimental section with a note on the computational performance of the CRO-SL algorithm proposed.

4.1. Data, algorithms for comparison and experimental parameters

In order to illustrate the performance of the chosen approach, we consider gridded wind speed monthly average data from the ERA-Interim reanalysis of the European Centre for Medium-Range Weather Forecasts (ECMWF) [50]. Data from January 1979 until July 2017 are taken, with a total of $\hat{t} = 462$ months available. We have divided them into training period ($t^T = 231$) months and test period ($t^V = 231$) months. Fig. 2 shows the location of the reanalysis nodes ($|S| = 540$ nodes considered). The wind power field is calculated by means of the standard approximation:

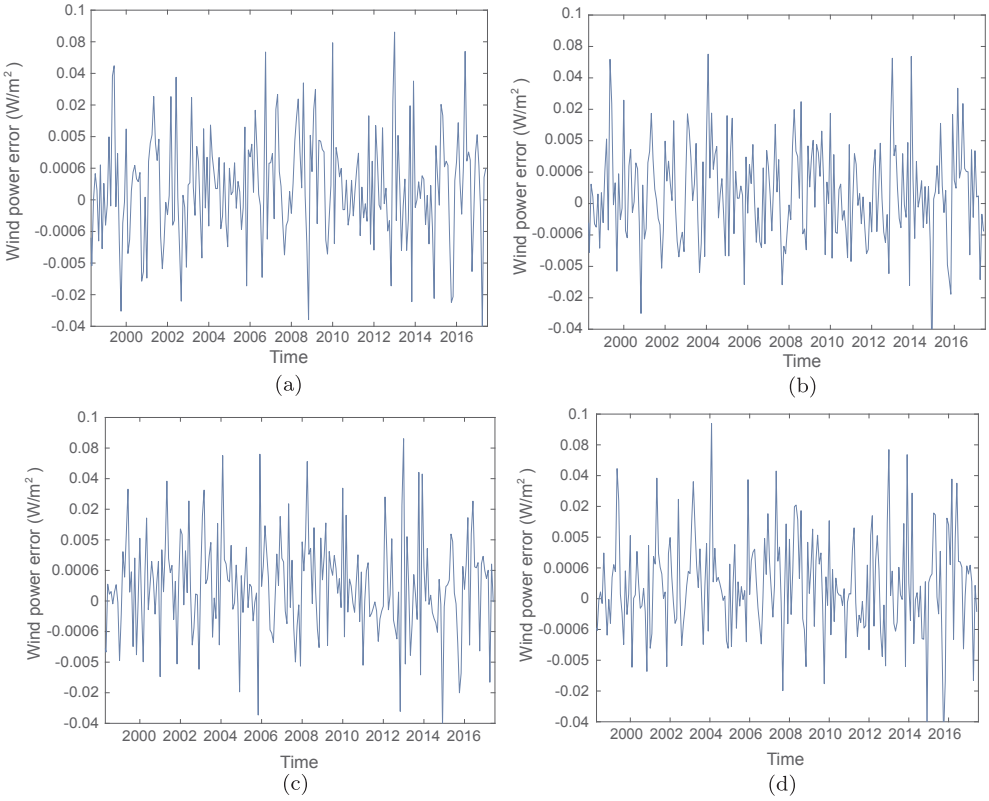

Fig. 6. Wind power reconstruction error (W/m^2) for the best measuring point reconstructed with the Analogue method, for different number of selected representative points; (a) $N = 5$; (b) $N = 10$; (c) $N = 15$ and (d) $N = 20$.

$$P = \frac{1}{2}\cdot\rho\cdot v^3 \tag{2}$$

where P stands for the wind power, in W/m^2, ρ is the dry wind density (1.225 kg/m^3 for air at 15 °C and at average sea level pressure), and v is the wind speed.

For comparison purposes, we consider a number of state-of-the-art meta-heuristic algorithms: a Harmony Search algorithm (HS) [48], an Evolutionary Algorithm (EA) [49] with two points crossover, tournament selection and uniform mutation, an integer version of the Particle Swarm Optimization approach (PSO) [51] and a Simulated Annealing algorithm (SA) [52]. Table 1 shows the parameters of the different approaches taken into account in all the experiments carried out in the paper.

4.2. Results and discussion

The experiments consist of running the CRO-SL and the alternative meta-heuristics for comparison in four different cases: $N = 5$, 10, 15 and 20. Note that they imply a reduction in the number of measuring points of 99%, 98%, 97% and 96%, respectively, to carry out the wind field reconstruction. Table 2 shows the results in terms of the RMSE in the wind power field reconstruction, with the CRO-SL and the meta-heuristics considered for comparison. As can be seen, the CRO-SL is able to obtain the best results consistently for all values of N. The EA obtains the second best solution in most cases, slightly better than HS, PSO and SA, respectively. The differences in performance are more acute when comparing the CRO-SL with the rest of meta-heuristics, which means that the search capabilities of the proposed approach are significant. Table 3 completes this information by showing a comparison of the best solution found with optimization algorithms (CRO-SL) against a randomly generated solution. Alternative statistical measures such as correlation coefficient (CC), bias (Bias), mean absolute error (MAE), normalized root mean square error (NRMSE) and scatter index (SI) are also reported.

Once we have spotted the CRO-SL as the best meta-heuristics among those tested, we apply it in order to study the RMSE dependence of N for large number of N values. The results are shown in Fig. 3. Note that each point corresponds to 5 CRO-SL runs (average value is displayed). Results for $N = 5$, 10, 15 and 20 are marked in red in the figure for reference. As can be seen, the RMSE obtained is smaller when N grows, as expected, until a value of around $N = 70$ stations, where it remains more or less constant. Note that the RMSE starts growing again for high values of N (from $N = 350$ stations), where it seems that the information is somehow degraded by the noise introduced when including a high number of measuring points in the wind field reconstruction.

Fig. 4 shows the best solutions obtained by the CRO-SL (measuring points selected), for the different cases considered. It is very interesting that the algorithm consistently selects points in the North Atlantic as those most informative for reconstructing the field of monthly wind power, mainly in the cases with lower N. When the number of possible measuring points grows ($N = 15$ and 20), the algorithm includes points over the Atlantic ocean, and additional a few points in the Mediterranean, as those with the most information to reconstruct the field. This result is clearly explained by the prevailing Westerlies in Europe, which make that points over the Atlantic ocean contain the most information to reconstruct the average monthly wind over the whole continent. Note that the best solution obtained does not include any in-land point. Instead, for every solution, the points are located over the ocean. This is again consistent, since there are no local influences introduced by orography and topography.

Fig. 5 shows the reconstruction error of the complete wind power field with the AM, for different values of N. The worst reconstruction is found in the North West Atlantic, and some parts of the Mediterranean Sea. This is consistent no matter the value of N considered to guide the reconstruction with the AM. The error range is in general good, with in-land values around 0.08 W/m^2 or even better, and ocean values around

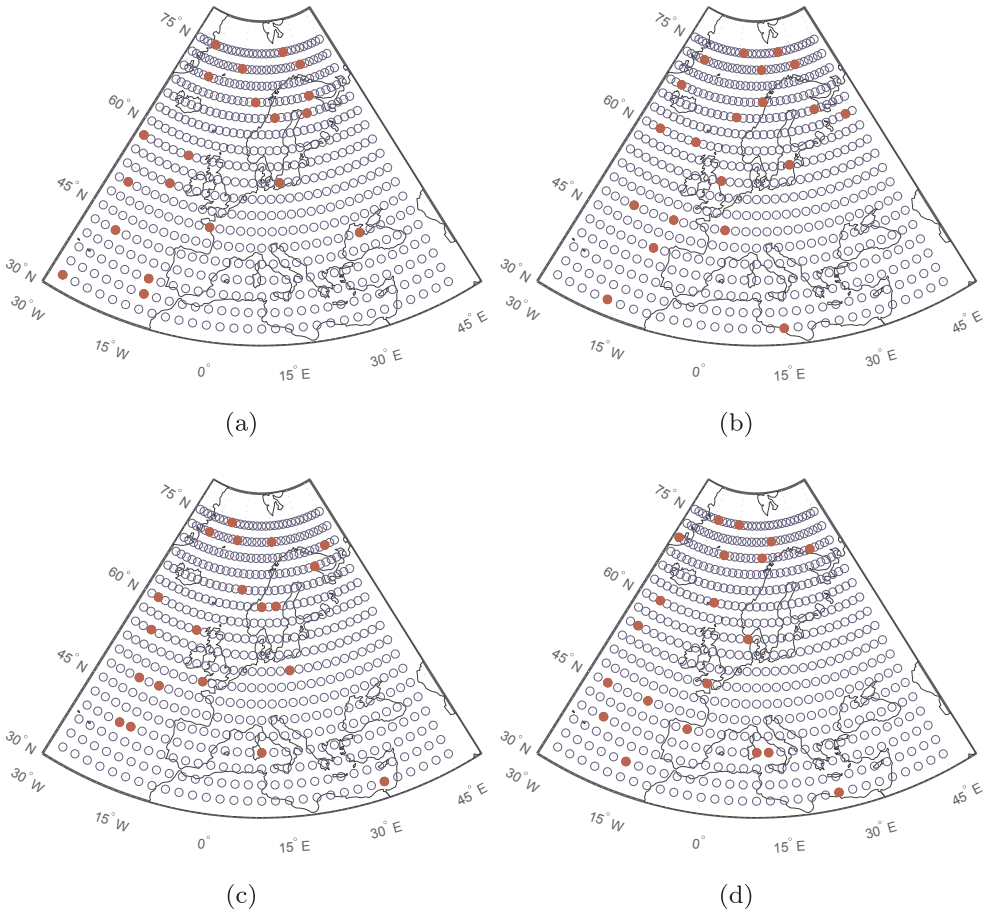

Fig. 7. Best solution found by the CRO-SL (20 representative points), in different seasons; (a) Spring; (b) Summer; (c) Autumn and (d) Winter.

0.6–0.9 W/m^2. In general the reconstruction RMSE for the complete wind power field is good, as can be see in Table 2, with values around 0.18–0.28 W/m^2 when the solutions by the CRO-SL are considered.

Fig. 6 shows the reconstruction error for the best point with the AM, in the test period. As can be seen, the error is small (less than 0.08 W/m^2 in all cases), and there are not big differences with $N = 5$ or $N = 20$. This indicates that the reconstruction method is strong even from a small number of measuring points to start with, so it is robust when dealing with scarce information for the reconstruction. Note that when the number of representative points grows, the algorithm tends to better covering some specific zones (the Atlantic Ocean), instead of selecting new ones. This is significant, since aims to highlight the importance of the selected zones for the reconstruction, discarding the selection of points in alternative areas.

Fig. 7(a), (b), (c) and (d) shows the best solutions obtained by the CRO-SL for the case $N = 20$, in different seasons: Spring, Summer, Autumn and Winter, respectively. It is possible to see how the representative points are again consistently located in the North Atlantic in all seasons. It seems that in Autumn and Winter the algorithm chooses some more points in the Mediterranean to characterize the wind power field than in Spring or Summer. In general the seasonal study shows similar zones as those obtained when all the data are managed.

Finally, we have identified the least representative points for the reconstruction, i.e., which set obtains the maximum error in the wind power field reconstruction. Fig. 8 shows the least representative points selected. As can be seen, in all cases the selected points are mainly located at the North of Scandinavia and some points in the Eastern Mediterranean. The former are located close to circumpolar circulation, local winds with deficit of information in the reconstruction of the whole wind power field. The latter could be related to the high seasonality of the circulation in that area, associated with the Etesian winds – local strong north winds of the Aegean Sea – which also seems to have little information for the reconstruction of the wind power field in Europe. This analysis shows the physical consistency of the results obtained by the CRO-SL with AM. Note that the proposed method does not consider any other climatological data source, but the wind power series in the past, which reflects the robustness of the method.

Finally, just a brief note on the specific CRO-SL performance. We can study it in terms of how the different substrates (search procedures) help the algorithm find better solutions. Fig. 9 shows the percentage of best larvae formed in each substrate, for the $N = 20$ case. It is easy to see that the 2Px and the MPx operators dominate the production of the best larvae, with the SM operator as a good contributor, specially in the first stages of the algorithm. The HS and DE substrates are those contributing the least to generate best larvae (solutions) in the algorithm's evolution. Another interesting analysis on the CRO-SL performance is given by Fig. 10, which shows the number of new larvae able to get into the reef during the CRO-SL evolution. In this case, the contribution of the 2Px and SM operators is higher than the rest in the first stages of the algorithm, while the 2Px and MPx finally dominate the inclusion of new larvae in the middle and final generations of the CRO-SL. Note that at the beginning of the algorithm the number of new larvae in the reef is significant, pushed by the fact that the number of holes in the reef is high, and the operators are able to find good quality solutions quickly. In the last stages of the algorithm, the number of new larvae in the reef is less important, and a minimum of larvae renewal due to depredation is maintained until the end of the algorithm.

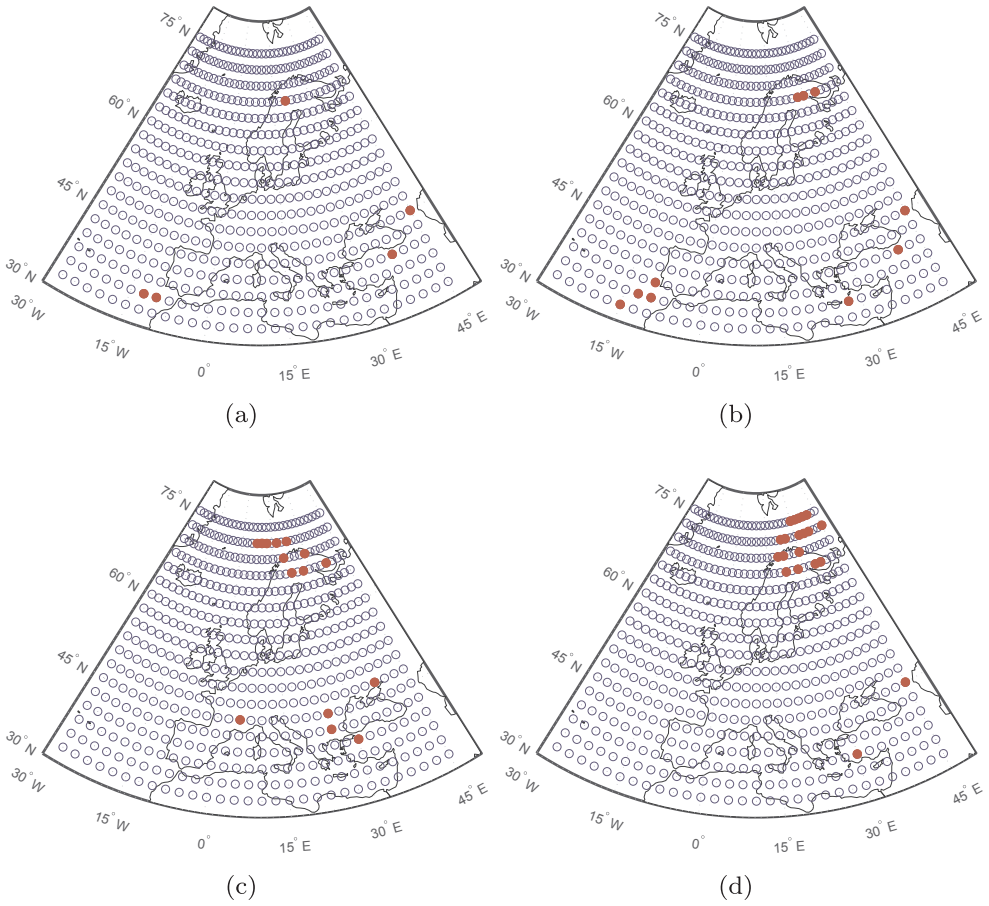

Fig. 8. Least representative points for the wind power field reconstruction; (a) $N = 5$; (b) $N = 10$; (c) $N = 15$ and (d) $N = 20$.

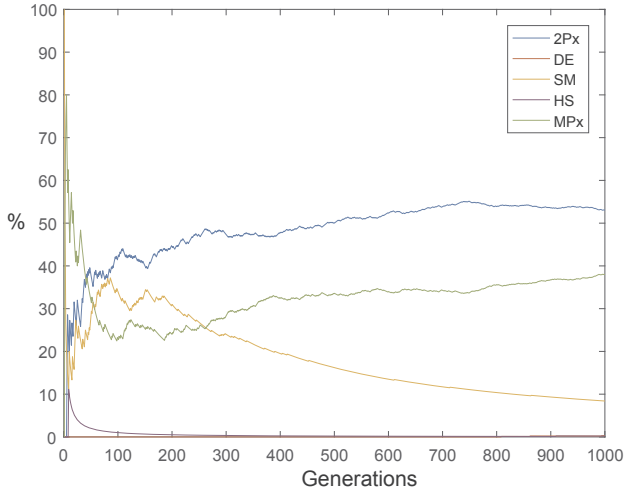

Fig. 9. Percentage of best larvae obtained in the CRO-SL from each substrate, in the case $N = 20$.

5. Conclusions

In this paper we have reconstructed monthly averaged wind power fields in Europe from a reduced number of representative points ($N = 5, \ldots, 20$) by applying a novel meta-heuristic algorithm, the Coral Reefs Optimization with Substrate Layer (CRO-SL). In this case, it has been combined with the Analogue Method (AM) as reconstruction procedure. An experimental evaluation of the method using the ERA-Interim reanalysis data (monthly resolution) has been carried out. The results reveal that the spatial pattern of the representative points is consistently located over the Atlantic ocean, which agrees with the predominance of the mid-latitude Westerlies in the region, whereas inland measuring points have less reconstructive information. The reconstruction of the complete wind power field with a reduced number of selected points is accurate, with average errors lower than 1 m/s (0.61 W/m^2) in the test period. The algorithm tends to use consistently points of the same areas as N increases. Thus, the performance improvement is obtained by increasing the sampling density in the key regions, rather than including new areas of information. An analysis of the least informative points has revealed that they are mainly located at the North of Scandinavia, and some points in the Eastern Mediterranean. The robustness of the proposed methodology has also been analyzed by performing an evaluation at seasonal scale, which has shown small differences in behavior among different seasons. Finally, note that this method can be applied in complete reconstructions of wind speed or power fields, useful for wind speed analysis and wind resource evaluation studies, but also to estimate the best location for measuring points or stations for long-term wind energy analysis. The proposed method also opens new lines of research to apply it to problems related to critical infrastructure selection, optimal selection of points for climate change control or long-term studies of alternative renewable-energy resource variables.

Acknowledgements

This work has been partially supported by the projects TIN2014-54583-C2-2-R, TIN2017-85887-C2-2-P and PALEOSTRAT (CGL2015-69699-R) of the Spanish Ministerial Commission of Science and Technology (MICYT), and by the Comunidad Autónoma de Madrid, under project number S2013ICE-2933_02.

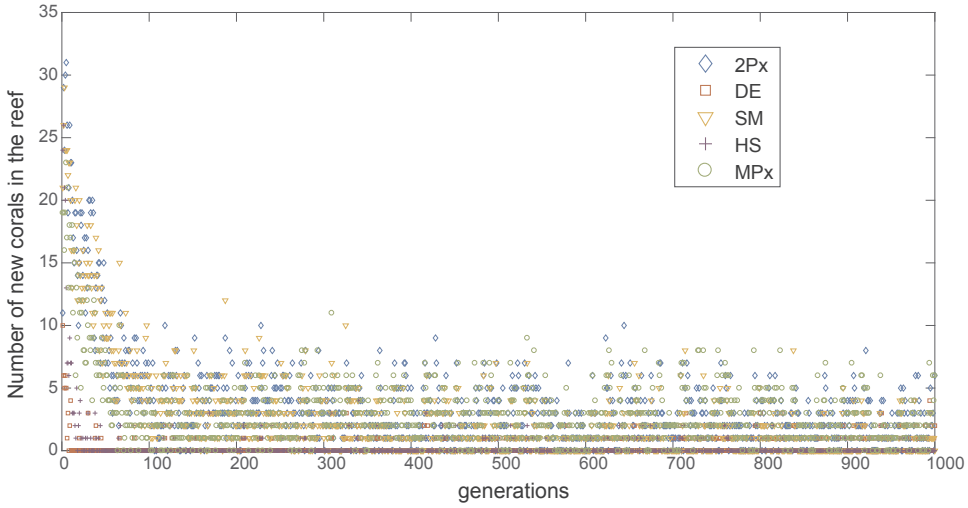

Fig. 10. Evolution of the number of new larvae in the CRO-SL which are able to get into the reef per generation and substrate, case $N = 20$.

References

[1] Ringkjob HK, Jourdier B, Drobinski P, Plougonven R, Tankov P. Modelling the variability of the wind energy resource on monthly and seasonal timescales. Renew Energy 2017;113:1434–46.

[2] Robert S, Foresti L, Kanevski M. Spatial prediction of monthly wind speeds in complex terrain with adaptive general regression neural networks. Int J Climatol 2013;33:1793–804.

[3] Tar K. Some statistical characteristics of monthly average wind speed at various heights. Renew Sustain Energy Rev 2008;12(6):1712–24.

[4] Kirchner-Bossi N, Prieto L, García-Herrera R, Carro-Calvo L, Salcedo-Sanz S. Multi-decadal variability in a centennial reconstruction of daily wind. Appl Energy 2013;105:30–46.

[5] Kirchner-Bossi N, García-Herrera R, Prieto L, Trigo RM. A long-term perspective of wind power output variability. Int J Climatol 2015;35(9):2635–46.

[6] Ahmed AS. Wind energy characteristics and wind park installation in Shark El-Ouinat, Egypt. Renew Sustain Energy Rev 2018;82:734–42.

[7] Souloukngа MH, Doka SY, Revanna N, Djongyang N, Kofane TC. Analysis of wind speed data and wind energy potential in Faya-Largeau, Chad, using Weibull distribution. Renew Energy 2018;121:1–8.

[8] Aquila G, Santana Peruchi R, Rotela Junior P, Souza Rocha LC, de Queiroz AR, de Oliveira Pamplona E, et al. Analysis of the wind average speed in different Brazilian states using the nested GR & R measurement system. Measurement 2018;115:217–22.

[9] Bianchi E, Solarte A, Guozden TM. Large scale climate drivers for wind resource in Southern South America. Renew Energy 2017;114:708–15.

[10] Pryor SC, Barthelmie RJ, Schoof JT. Inter-annual variability of wind indices across Europe. Wind Energy 2006;9:27–38.

[11] Cannon DJ, Brayshaw DJ, Methven J, Coker PJ, Lenaghan D. Using reanalysis data to quantify extreme wind power generation statistics: a 33 year case study in Great Britain. Renew Energy 2015;75:767–78.

[12] Pryor SC, Barthelmie RJ. Climate change impacts on wind energy: a review. Renew Sustain Energy Rev 2010;14:430–7.

[13] Pryor SC, Schoof JT, Barthelmie RJ. Climate change impacts on wind speeds and wind energy density in northern Europe: empirical downscaling of multiple AOGCMs. Clim Res 2005;29:183–98.

[14] François B, Martino S, Tøfte LS, Hingray B, Mo B, Creutin JD. Effects of increased wind power generation on mid-Norway's energy balance under climate change: a market based approach. Energies 2017;10(227):1–18.

[15] Pryor SC, Barthelmie RJ. Assessing climate change impacts on the near-term stability of the wind energy resource over the United States. Proc Natl Acad Sci 2011;108(20):8167–71.

[16] Zheng CW, Li CY, Li X. Recent decadal trend in the North Atlantic wind energy resources. Adv Meteorol 2017;2017:1–8. id. 7257492.

[17] Zheng CW, Xiao ZN, Zhou W, Chan XB, Chen X. 21st Century maritime silk road: a peaceful way forward. Springer; 2018.

[18] Kwon SD. Uncertainty analysis of wind energy potential assessment. Appl Energy 2010;87(3):856–65.

[19] Jung S, Vanli OA, Kwon SD. Wind energy potential assessment considering the uncertainties due to limited data. Appl Energy 2013;102:1492–503.

[20] Amirinia G, Kamranzad B, Mafi S. Wind and wave energy potential in southern Caspian Sea using uncertainty analysis. Energy 2017;120:332–45.

[21] Amirinia G, Mafi S, Mazaheri S. Offshore wind resource assessment of Persian Gulf using uncertainty analysis and GIS. Renew Energy 2017;113:915–29.

[22] Saavedra-Moreno B, Salcedo-Sanz S, Carro-Calvo L, Gascón-Moreno J, Jiménez-Fernández S, Prieto L. Very fast training neural-computation techniques for real measure-correlate-predict wind operations in wind farms. J Wind Eng Ind Aerodyn 2013;116:49–60.

[23] Díaz S, Carta JA, Matías JM. Performance assessment of five MCP models proposed for the estimation of long-term wind turbine power outputs at a target site using three machine learning techniques. Appl Energy 2018;209:455–77.

[24] Ritter M, Deckert L. Site assessment, turbine selection, and local feed-in tariffs through the wind energy index. Appl Energy 2017;185(Part 2):1087–99.

[25] Wang H, Kawahara Y, Weng C, Yuan J. Representative selection with structured sparsity. Pattern Recogn 2017;63:268–78.

[26] Rife SL, Vanvyve E, Pinto JO, Monaghan AJ, Davis CA, Poulos GS. Selecting representative days for more efficient dynamical climate downscaling: application to wind energy. J Appl Meteorol Climatol 2013;52:47–63.

[27] Ruane AC, McDermid SP. Selection of a representative subset of global climate models that captures the profile of regional changes for integrated climate impacts assessment. Earth Perspect 2017;4(1):1–20.

[28] Wilcke RA, Bärring L. Selecting regional climate scenarios for impact modelling studies. Environ Modell Softw 2016;78:191–201.

[29] Lutz F, ter Maat HW, Biemans H, Shrestha AB, Westerdan P, Immerzeel WW. Selecting representative climate models for climate change impact studies: an advanced envelope-based selection approach. Int J Climatol 2016;36:3988–4005.

[30] Amorim AM, Gonçalves AB, Miguel Nunes L, Sousa AJ. Optimizing the location of weather monitoring stations using estimation uncertainty. Int J Climatol 2012;32:941–52.

[31] Salcedo-Sanz S, Pastor-Sánchez A, Blanco-Aguilera A, Prieto L, García-Herrera R. Feature selection in wind speed prediction systems based on a hybrid coral reefs optimization – extreme learning machine approach. Energy Convers Manage 2014;87:10–8.

[32] Salcedo-Sanz S, Jiménez-Fernández S, Aybar-Ruiz A, Casanova-Mateo C, Sanz-Justo J, García-Herrera R. A CRO-Species optimization scheme for robust global solar radiation statistical downscaling. Renew Energy 2017;111:63–76.

[33] Salcedo-Sanz S, Deo RC, Cornejo-Bueno L, Camacho-Gómez C, Ghimire S. An efficient neuro-evolutionary hybrid modelling mechanism for the estimation of daily global solar radiation in Sunshine State of Australia. Appl Energy 2018;209:79–94.

[34] Lorentz E. Atmospheric predictability as revealed by naturally occurring analogues. J Atmos Sci 1969;26:636–46.

[35] Alessandrini S, Delle Monache L, Sperati S, Cervone G. An analog ensemble for short-term probabilistic solar power forecast. Appl Energy 2015;157:95–110.

[36] Alessandrini S, Delle Monache L, Sperati S, Nissen JN. A novel application of an analog ensemble for short-term wind power forecasting. Renew Energy 2015;76:768–81.

[37] Vanvyve E, Delle Monache L, Monaghan AJ, Pinto JO. Wind resource estimates with an analog ensemble approach. Renew Energy 2015;74:761–73.

[38] Horton P, Jaboyedoff M, Obled C. Global optimization of an Analog Method by means of Genetic Algorithms. Mon Weather Rev 2017;145:1275–94.

[39] Salcedo-Sanz S, del Ser J, Landa-Torres I, Gil-López S, Portilla-Figueras A. The Coral Reefs Optimization algorithm: a novel metaheuristic for efficiently solving optimization problems. Sci World J 2014.

[40] Salcedo-Sanz S. A review on the coral reefs optimization algorithm: new development lines and current applications. Prog Artif Intell 2017;6:1–15.

[41] Salcedo-Sanz S, Muñoz-Bulnes J, Vermeij M. New coral reefs-based approaches for the model type selection problem: a novel method to predict a nation's future energy demand. Int J Bio-Inspired Comput 2017;10(3):145–58.

[42] Salcedo-Sanz S, Camacho-Gómez C, Molina D, Herrera F. A Coral Reefs Optimization algorithm with substrate layers and local search for large scale global optimization. In: Proc. of the IEEE world congress on computational intelligence, Vancouver, Canada; July, 2016.

[43] Salcedo-Sanz S, Camacho-Gómez C, Mallol-Poyato R, Jiménez-Fernández S, del Ser J. A novel Coral Reefs Optimization algorithm with substrate layers for optimal battery scheduling optimization in micro-grids. Soft Comput 2016;20:4287–300.

[44] Salcedo-Sanz S, Camacho-Gómez C, Magdaleno A, Pereira E, Lorenzana A.

Structures vibration control via tuned mass dampers using a co-evolution coral reefs optimization algorithm. J Sound Vib 2017;393:62–75.
[45] Camacho-Gómez C, Wang X, Pereira E, Díaz IM, Salcedo-Sanz S. Active vibration control design using the Coral Reefs Optimization with Substrate Layer algorithm. Comput Struct 2018;157:14–26.
[46] Salcedo-Sanz S. Modern meta-heuristics based on nonlinear physics processes: a review of models and design procedures. Phys Rep 2016;655:1–70.
[47] Storn R, Price K. Differential Evolution – a simple and efficient heuristic for global optimization over continuous spaces. J Global Optim 1997;11:341–59.
[48] Geem ZW, Kim JH, Loganathan GV. A new heuristic optimization algorithm: harmony search. Simulation 2001;76(2):60–8.
[49] Eiben AE, Smith JE. Introduction to evolutionary computing. 1st ed. Natural computing series Springer-Verlag; 2003.
[50] Dee DP, M Uppala S, Simmons AJ, Berrisford P, Poli P, Kobayashi S, et al. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Quart J Roy Meteorol Soc 2011;137:553–97.
[51] Kennedy J, Eberhart R. Particle swarm optimization. In: Proc. of the 4th IEEE international conference on neural networks; 1995. p. 1942–8.
[52] Kirpatrick D, Gerlatt CD, Vecchi MP. Optimization by simulated annealing. Science 1983;220:671–80.